Vision Research 50 (2010) 904–913

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Object speed derived from the integration of motion in the image plane and motion-in-depth signaled by stereomotion and looming

Sieu K. Khuu [a,*], Terence C.P. Lee [b], Anthony Hayes [c,d]

[a] The School of Optometry and Vision Science, The University of New South Wales, Sydney 2052, Australia
[b] Psychology Unit, Hong Kong Baptist University, Shek Mun, Shatin, Hong Kong SAR, China
[c] School of Psychology, University College Dublin, Belfield, Dublin 4, Ireland
[d] Division of Psychology, Nanyang Technological University, Singapore 639798

ARTICLE INFO

Article history:
Received 24 May 2009
Received in revised form 6 February 2010

Keywords:
Speed processing
Motion-in-depth
Stereomotion
Looming

ABSTRACT

We investigate the influence of local motion in the retinal image plane on the perception of speed-in-depth. Observers judged the apparent speed-in-depth of a square plane of dynamic dots that moved towards the observer. Dots forming the surface of the plane underwent random-direction motion in the image plane. We examined the consequences of changing the dots' image-plane speed on the apparent speed of the stimulus as it traversed depth, where depth is signaled by stereomotion or looming. Results for both the stereomotion and looming conditions indicate that as the speed of random-direction motion in the image plane increases, the apparent speed-in-depth of the stimulus also increases. When stereomotion was used to signal motion-in-depth, the speed judgment is adequately modeled by the resultant of a vector sum of dot-speed in the image plane and speed-in-depth. However, when looming was used to define motion-in-depth, a different pattern of results was found – the apparent speed-in-depth is lower than the actual speed-in-depth, and the results are best predicted by simple averaging. Our results demonstrate that the integration of speed in the image plane and speed-in-depth, to determine object speed, is dependent on the type of cue used to signal motion-in-depth, and this difference is a consequence of the ways in which looming and stereomotion cue motion-in-depth. Looming is derived not at a local stage of motion analysis, but is available only via global integration of local velocities, and consequently global speed conforms to the average speed. Stereomotion, on the other hand, provides an effective cue for individuating local velocities in depth.

1. Introduction

The visual world is usually dynamic, consisting of objects and local textures that are constantly in motion. To successfully interact with this dynamic environment, the visual system must accurately estimate the motion of objects. In particular, neural mechanisms responsible for the extraction of motion are designed to recover two key properties of image motion: speed (which specifies the rate of motion) and direction (which indicates the trajectory of motion). Under most circumstances this recovery is performed with a high degree of efficiency and accuracy. For example, experiments consistently reveal the capability of the human visual system to discriminate minute differences for both speed and direction (e.g., Ledgeway, 1999; McKee, 1981).

Determination of the motion of objects in natural scenes is not a straightforward computation; objects move in three-dimensional space, and any object velocity can be encoded as motion signals present in the two dimensions of the image plane – the retinal projection plane – and in the orthogonal dimension of depth. Adding to the complexity, motion-in-depth is not explicitly encoded via a projection surface of the human visual system in the third dimension, but must be inferred from a host of visual cues to depth. While the neural mechanisms responsible for the extraction of motion in the image plane are likely to be spatio-temporal operators, with image motion registered by extended integration of detectors that sample information at successive positions on the retinal image (Reichardt, 1961), the extraction of motion-in-depth is thought to be facilitated by cues such as binocular disparity, looming, and occlusion, which serve to distinguish relative location in depth (see, e.g., Howard & Rogers, 1995; Poggio, Gonzalez, & Krause, 1988). The coding of image-plane motion, rendered on a projection surface, and the coding of motion-in-depth, on a non-projection "surface", very likely accounts for the functional differences found between the extraction of motion in the image plane and motion-in-depth. Given that objects in the world usually simultaneously traverse both the image plane and depth, under most circumstances the visual system is likely to consider the mutual processing of motion signals in the image plane and in depth when determining object motion. However, the nature and extent of this

* Corresponding author.
E-mail address: s.khuu@unsw.edu.au (S.K. Khuu).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.02.005

interaction is at present unclear, particularly for the perception of image speed.

Insight into the interaction of speed in the image plane and speed-in-depth comes from studies that report an illusory difference in apparent speed between radial and rotary, complex, global dot-motion (GDM) patterns. With this stimulus, a percept of coherent motion is derived solely by the coherent movement in the image plane of a number of dots along radial trajectories confined to a fixed spatial area. Clifford, Beardsley, and Vaina (1999) reported that radial expanding GDM patterns are perceived to be moving faster than rotary patterns with the same dot-speed distribution. This speed effect is also observed with stimuli consisting of a small number of image-plane gratings oriented to simulate radial motion (Bex & Makous, 1997). A likely account for this speed illusion stems from the discrepancy between a perceived velocity in depth and the actual image-plane estimate of velocity on the surface of the retina. When depth trajectories are projected onto an image-plane surface (such as the retina), displacement in the image plane will often be smaller than the displacement in depth that produced it. Thus, radial motion might appear to be faster than rotary motion if the visual system extracts a speed estimate based on the possible trajectory in depth, rather than the motion in the image plane. These findings are important in revealing "a priori" assumptions made in the interpretation of complex global motion as signaling motion-in-depth, when the motion pattern is in an appropriate configuration.

The findings of the studies briefly described above indicate a close relationship between the perception of speed in the image plane and speed-in-depth, and suggest that there is interaction between the two. Evidence in support of the notion of this interaction in the perception of speed comes from studies that examined the relationship between stereomotion and looming (stereomotion is generated by a systematic change in the disparity of an object over time; looming is essentially an image-plane signal – a gradual expansion or magnification of image size). For example, Gray and Regan (1998) showed that speed discrimination thresholds are lower for stimuli in which looming and stereomotion are both present, than when these cues are presented individually. Moreover, a procedure that encodes and sums the speed from looming and disparity can provide an effective account for human performance (Hogervorst & Brenner, 2004).

While the findings of Gray and Regan (1998) and Hogervorst and Brenner (2004) implicate an interactive relationship between the perception of looming and speed-in-depth (from stereomotion), a number of issues remain unresolved regarding the nature of this interaction. Looming is derived from local motion on the image plane, but as the findings of Gray and Regan (1998), Hogervorst and Brenner (2004), and the observations made in investigations of optic flow, make clear, the interpretation of a radially expanding stimulus – a looming stimulus – is one of motion-in-depth. Thus, in the perception of looming, the visual system is not sensitive to the actual speed on the image plane of local texture, but instead "infers" the motion-in-depth from the velocities on the image plane. This is problematic because these cases of interactions between looming and stereomotion do not provide a reflection of the ability of the visual system to combine "actual" speed in the image plane (movement that is not inferred as motion-in-depth) and speed-in-depth. It remains unclear from available data whether any interaction occurs when determining motion in the different dimensions when deriving image speed. Indeed, motion in the image plane in natural scenes often does correlate with motion-in-depth, and it is possible that speed signals in the image plane that are uncorrelated with depth may act as spurious noise to the computation of speed-in-depth, and under these circumstances the visual system treats these speed estimates independently. Such a situation would imply that separate mechanisms exist for the extraction of speed in the image plane and speed-in-depth.

Motion-in-depth can be signaled by a variety of depth cues, of which stereomotion and looming are particularly noteworthy. To what extent speed in the image plane affects the perception of motion-in-depth signaled by stereomotion and by looming, and whether the computation is similar for both these cues, are important questions. Published research has reported a strong interactive link between these two depth cues, perhaps stemming from the fact that in natural scenes they concurrently define the motion-in-depth of objects. It is possible that the same computation is used to derive speed signaled both by looming and by stereomotion, and published findings provide evidence of mutual processing of stereomotion and looming. Regan and Beverley (1979) found that both the sensation of motion-in-depth and a motion-in-depth aftereffect, signaled by either stereomotion or looming, can be canceled by the other cue, suggesting that both cues feed into a common stage of motion-in-depth processing. Furthermore, when both cues are available, greater accuracy is found for estimations of time to collision with approaching objects (Gray & Regan, 1998; Heuer, 1993), of timing of ball-catching movements (Rushton & Wann, 1999; Savelsbergh, Whiting, & Bootsma, 1991), and of perceived speed and direction of motion-in-depth embedded in optic flow (Gray, Macuga, & Regan, 2004). These findings can be explained by a model that combines the two cues as a weighted sum, in which the weighting is adjusted according to the reliability of each cue (Regan & Beverley, 1979; Regan & Gray, 2000; Rushton & Wann, 1999).

Despite the evidence of early cue integration, the mechanisms underlying stereomotion and looming display many differences in their functional properties, which suggests there is no simple common interactive relationship between the two cues. Both from theoretical inference and empirical evidence, Regan and Beverley (1979) argued that stereomotion is more effective as a cue to motion-in-depth for small objects at high speed, whereas looming is more effective for large objects at low speed. Results consistent with these findings were also obtained for time-to-collision discriminations and direction estimations of approaching objects (Gray & Regan, 1998; Heuer, 1993). Moreover, observers who have selective stereomotion blindness have been reported in the literature (Richards, 1977; Richards & Regan, 1973), but there is no evidence for the analogous case of looming-blind observers (Regan & Beverley, 1983). Additionally, the presence of reference marks is necessary to perceive motion-in-depth signaled by stereomotion, but not for looming (Regan, Erkelens, & Collewijn, 1986). Also, the detection of stereomotion and the detection of looming follow different computational paths: while stereomotion can be extracted at the local level of motion processing, perhaps by disparity-tuned neurons in areas early in the processing stream, looming must be extracted at a global stage where local velocities are integrated to determine the rate of expansion. These functional differences between looming and stereomotion provide grounds to question whether the same speed computation is applied to each in generating an estimate of speed-in-depth that is derived from the two cues.

The purpose of the present study is to examine the degree to which speed in the image plane and speed-in-depth are combined. We investigated the conditions under which, and the nature of, the contribution of "uncorrelated" speed in the image plane to the percept of speed-in-depth cued both by stereomotion and by looming. In a related study, Khuu and Badcock (2002) examined the extent to which the apparent speed of a moving cloud of dots is affected by the local speed of constituent dots undergoing random motion. They reported that as the local speed of dots increased so did the apparent speed of the moving cloud of dots, and the rate of change is consistent with a simple averaging between the object speed and the local dot-speed. In the present study we adopt an analogous stimulus to Khuu and Badcock (2002), a moving cloud of dots in the image plane, but our dot-defined stimulus also underwent motion-in-depth. Our stimulus was a square plane of dots that traversed depth signaled by either a change in binocular disparity (stereomotion), or by a systematic change in image size (looming), and, simultaneously, the same group of dots underwent random-direction motion in the image plane (see, Fig. 1A). This stimulus affords the advantage of allowing independent variation of the speed of dots in the image plane and the speed-in-depth of the dots. It is important to note that there is no perceptual confusion of the random motion signals in the image plane introduced in this stimulus as constituents of motion-in-depth, since it generates a clear percept of a swarming surface moving along the axis orthogonal to the image plane. In Experiment 1, we performed a matching experiment to ensure that the speed in the image plane and the speed-in-depth employed in the study were perceptually equal. Using these matched speeds, in Experiment 2 we examined the extent to which systematically changing the dot speed in the image plane affects the perception of the entire plane of dots traversing in depth, cued either by stereomotion or by looming.

2. Experiment 1: matching image-plane speed and speed-in-depth

Brenner, van den Berg, and van Damme (1996) reported that the motion-in-depth of a rigid object, signaled only by changing disparity, is perceptually slower than the same object that undergoes motion in the image plane at the same physical speed. This perceptual difference is problematic for the present study since an accurate analysis of the combination of speed in the image plane and speed-in-depth requires that both quantities are perceptually equal. To effectively examine the relative contribution of motion in the image plane and motion-in-depth in determining object speed, the extent to which they differ must be quantified for the range of speeds employed, and it must be individuated for different observers. In Experiment 1, we measured this characteristic of motion-in-depth using Method of Adjustment in which observers were required to equate the speed of two square planes of dots presented sequentially. One square moved towards the observer, while the other underwent motion to the right, exclusively in the image plane (Fig. 1A). In Experiment 1, similar to Brenner et al. (1996), the object's motion-in-depth was signaled by stereomotion (a systematic change in disparity), but in our experiment we additionally verified whether there was a perceptual speed difference between motion in the image plane and motion-in-depth signaled by looming (a systematic change in image size).

2.1. Method

2.1.1. Observers

Five observers participated in the stereomotion condition, while seven observers, including the five in the stereomotion condition, participated in the looming condition. All had normal or corrected-to-normal visual acuity. SKK and TL were authors; the other observers were experienced psychophysics observers who were naïve to the objectives of the experiment.

2.1.2. Stimulus

Stimuli were constructed from two monocular images, paired to construct a stereogram in which monocular images are viewed separately by both eyes, containing an orthographically presented

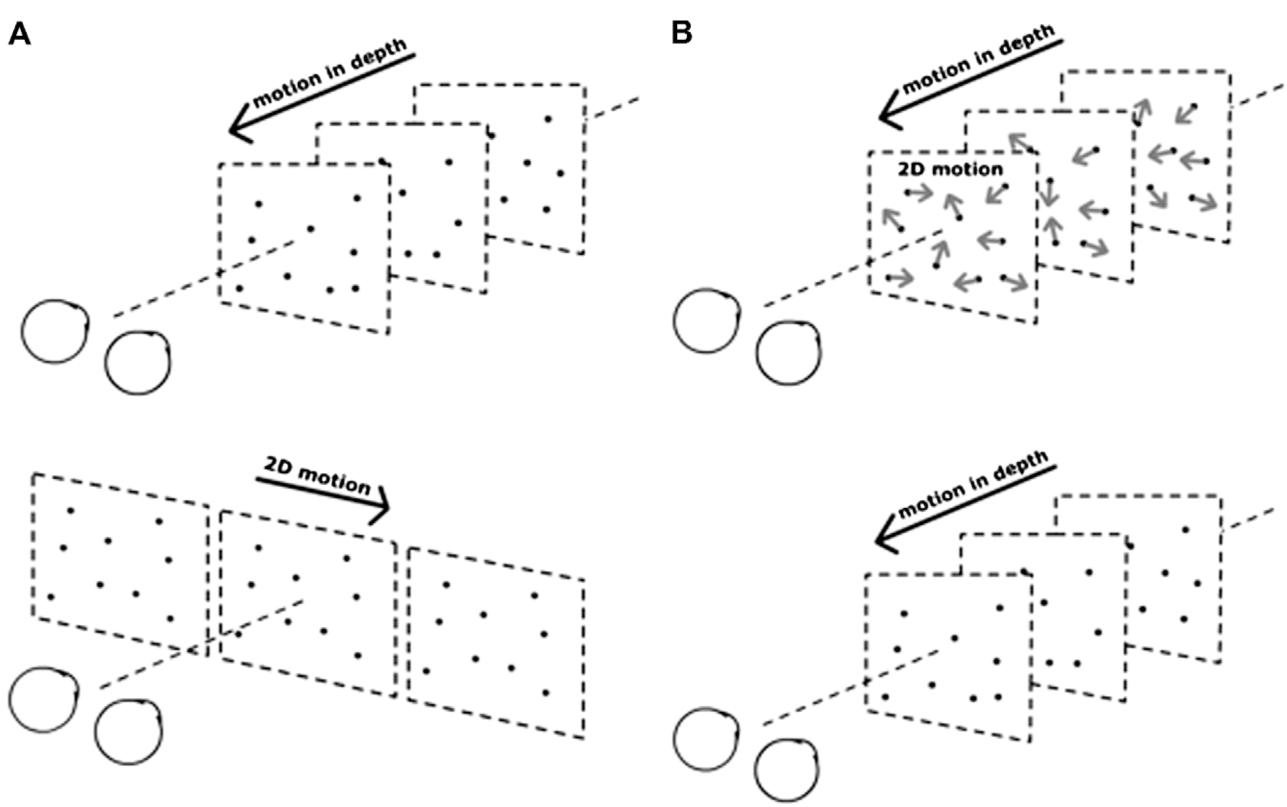

Fig. 1. Experimental design. (A) In Experiment 1, a dot-defined square was moved either towards the observer, or across the two-dimensional image plane. Either looming or stereomotion specified motion-in-depth. Observers were required to match the apparent speed in the two orthogonal directions. (B) In Experiment 2, the dot-defined square was moved towards the observer in each of two intervals. In one interval random-direction motion in the image plane was incorporated into the stimulus. Observers were required to match the apparent speed-in-depth of the two motion sequences.

square ($4.63^\circ \times 4.63^\circ$ at zero disparity) defined by 100 white circular anti-aliased dots (132 cd/m^2; diameter 0.11°) that occupied randomly chosen positions on a gray background (30 cd/m^2). To prevent tracking of individual dots, all dots were generated asynchronously with a limited lifetime of 133.33 ms, and were replotted back into the square upon expiry. One stimulus consisted of a stereogram containing dots undergoing rigid motion to the right, while the other stimulus consisted of a stereogram in which dots underwent motion-in-depth, as cued by stereomotion or by looming.

2.1.3. Motion in the image plane

To generate sequences of translational motion, all dots in each monocular image moved rightwards, giving rise to the impression of a dot-defined square translating across the screen. The speed of motion in the image plane was varied from trial to trial, subject to adjustments made by observers (see Procedure).

2.1.4. Stereomotion stimuli

Stereomotion was generated by moving corresponding dots in each monocular image towards each other horizontally at three monocular speeds (0.44 deg/s, 0.67 deg/s, or 1.00 deg/s) to simulate three levels of speed-in-depth. Consequently, the square appeared to move in depth towards the observer. By simple geometry, the arrays of speed in the image plane and speed-in-depth correspond to identical speed levels of 4.01 cm/s, 6.02 cm/s, or 9.03 cm/s. The disparity values were translated into depth values by the equation, $\eta \approx I\ \delta/D^2$; where η is the disparity, I is the interocular separation, δ is the simulated depth, and D is the viewing distance. The speed-in-depth, V_z, was derived from the equation, $V_z \approx D^2\ V/I$; where V is the rate of change of disparity.

2.1.5. Looming stimuli

To simulate looming, dots defining the square moved away from the center of the display to simulate an expanding square. To correlate with the motion specified by stereomotion, the dot coordinates were determined by the equation, $f(d') = q(d + d')/d$; where d is the viewing distance, d' is the simulated depth from disparity, q is the $\times$ or y coordinate with respect to the center of the display, and $f(d')$ is the re-scaled coordinate. The three levels of speed-in-depth (4.01 cm/s, 6.02 cm/s, or 9.03 cm/s) were simulated by dots moving along radial trajectories from the center of the stimulus at three ratios of expansion, 1.10, 1.15, or 1.22. Individual dots did not undergo a change in size. To simulate a randomized distance-in-depth traversed by the square, its range of angular substense was randomized from between 4.48° and 4.78° to between 4.33° and 4.93° for every motion sequence.

It is important to note that using the above calculations the speeds in the image plane of dots producing looming are not uniform within the stimulus, but are dependent on two factors. First, the speed of any dot is dependent on its position relative to the center of the stimulus, with dots further away moving at faster speeds. Second, the speed of dots is dependent on the simulated distance-in-depth from the observer. The nearer to the observer the object appeared, the greater is the looming and thus the faster the speed in the image plane. In Experiment 1, the distance-in-depth traversed by the square stimulus was randomized between 2.68 cm and 5.35 cm, and was symmetrical about the fixation plane (see Procedure), and to produce speeds in depth of 4.01 cm/s, 6.02 cm/s, or 9.03 cm/s, the fastest dot speeds in the image plane corresponding to the edge of the looming dot square were 0.096 deg/s, 0.384 deg/s, and 0.448 deg/s.

All stimuli were generated on a Macintosh G4 1.3 GHz computer using custom software written in MATLAB (version 5.3), and were displayed on a linearised monitor with a refresh rate of 100 Hz. A custom-built Wheatstone stereoscope was used to present stimuli at a viewing distance of 41 cm. To minimize eye movements, a fixation cross was presented at the center of the stimulus.

2.2. Procedure

Two of the aforementioned motion sequences were presented in random order with a 500 ms interval in between sequences (see Fig. 1). In one sequence, the dot-defined square underwent motion in the image plane (the test stimulus), but moved in depth (defined by stereomotion or looming; the reference stimulus) in the other sequence (see Fig. 1A). The task of the observer was to adjust the physical speed of dots in the image plane of the test stimulus to match the speed-in-depth of the reference stimulus. In a trial, the observer viewed a number of discrete presentations of the test and reference stimulus (in sequential order). At the start of each trial and on the first presentation, the image-plane speed of the test stimulus was randomized between 0.5 cm/s and 8 cm/s, while speed-in-depth of the reference stimulus was fixed to one of three different speeds, 4.01 cm/s, 6.02 cm/s, or 9.03 cm/s. Additionally, across different presentations within a trial, the distance traversed by the square moving in depth was randomized between 2.68 cm and 5.35 cm (thus duration was speed dependent), and was symmetrical about the fixation plane. This procedure ensured that observers could not reliably judge speed-in-depth by relying on the distance traveled in depth by the stimulus. After the first and for subsequent presentations (of the test and reference stimuli), the observer pressed keys to increase or decrease speed in the image plane by a constant step size of 0.14 deg/s (equivalent of 0.1 cm/s in depth). After this adjustment was made, a new presentation followed in which the previously adjusted speed was used to define the speed of the test stimulus. The observer repeated this task until he/she ended a trial when a perceptual match was achieved between the test and reference stimuli. Throughout the experiment, observers were required to focus on a fixation cross ($0.26^\circ \times 0.26^\circ$) at the center of the display at zero disparity. A block of trials consisted of six trials in which speed adjustments were made for each of the three speed-in-depth conditions and for both looming and steremotion, in random order. Observers each completed 10 blocks such that each condition had 10 trials. Results were averaged over the 10 trials for each condition.

2.3. Results and discussion

Fig. 2 plots the speeds of the square undergoing motion in the image plane matched to a square traversing depth at speeds of 4.01 cm/s, 6.02 cm/s, or 9.03 cm/s, for different observers. The dotted line represents veridical matches between speed in the image plane and speed-in-depth. There are two main findings. First, the pattern of results is similar for all observers (with some individual variability, perhaps due to the subjective nature of the task) and motion-in-depth appeared to observers to be approximately 1.5–3 times slower than motion in the image plane. Observers decreased the physical speed of the test stimulus (undergoing translational motion in the image plane) to perceptually match speed-in-depth. In other words, there is an underestimation of speed-in-depth when compared to speed in the image plane. Second, this data trend, consistent with Brenner et al. (1996), is evident for stereomotion, but is also present for looming despite the fact that looming is essentially a monocular cue to motion-in-depth. This finding clearly suggests that the visual system does not rely on the local speeds in the image plane signaled by individual dots, but instead integrates local velocities to extract looming, and given the similarity with stereomotion, interprets this stimulus as motion-in-depth. This finding reinforces the notion that looming, whilst an image-plane stimulus, is interpreted by the visual system

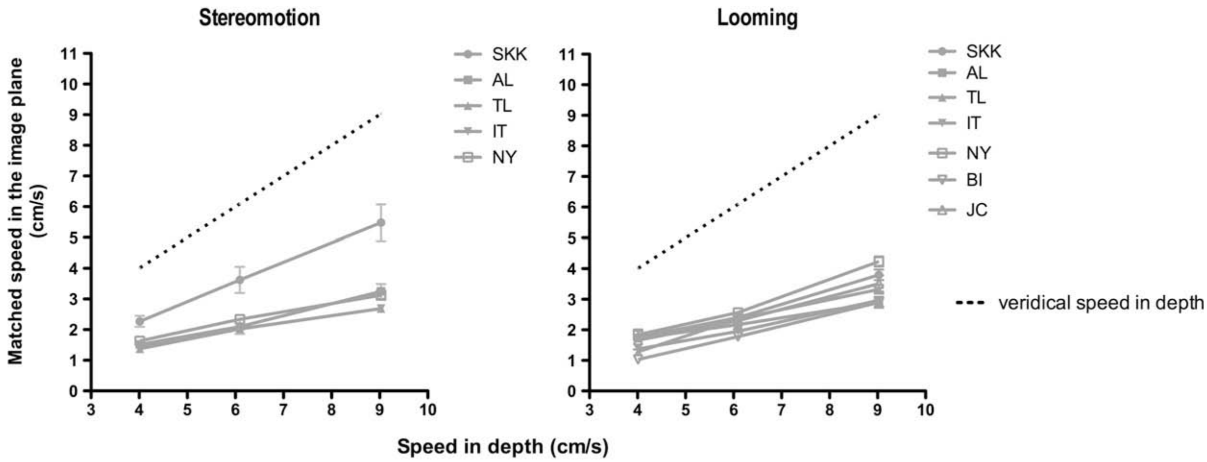

Fig. 2. The matched speeds in the image plane of the square undergoing translational motion are plotted against the speed of the comparison square undergoing motion-in-depth. The dotted line represents veridical matches between speed in the image plane and speed-in-depth. Error bars represent ±1 SEM.

as a strong cue to motion-in-depth, and is comparable to stereomotion.

The similarity in judging speed-in-depth signaled by stereomotion and by looming is perhaps indicative of analogous processing, and psychophysical evidence and computational analysis have suggested that the visual system readily combines motion-in-depth signals derived from stereomotion and from looming (Hogervorst & Brenner, 2004; Regan & Beverley, 1979; Regan & Gray, 2000; Rushton & Wann, 1999). This finding also validates the manner in which we generate our looming stimulus, which was simulated with dots of a constant size where this "cue-conflict" may potentially reduce the effectiveness of looming. However, the similarity between the results from looming and stereomotion demonstrates the effectiveness of dot-motion generated looming as a cue for signaling motion-in-depth.

It is important to note that the results of Experiment 1 are inconsistent with the observer relying on the monocular speed present in both monocular images (forming the stereogram) as the basis for perceiving speed-in-depth. As mentioned above, for stereomotion and looming conditions to generate apparent speeds in depth of 4.01 cm/s, 6.02 cm/s, or 9.03 cm/s, dots in the monocular image moved at maximum speeds of 0.44 deg/s, 0.67 deg/s, or 1.00 deg/s (equivalent to 0.314 cm/s, 0.47 cm/s, 0.715 cm/s) and 0.096, 0.384, 0.448 deg/s (equivalent to 0.07 cm/s, 0.27 cm/s, and 0.32 cm/s), respectively. Thus, if the visual system were to rely on the image-plane speed present in the monocular image, the matched speeds in the image plane of the square would correspond to these speeds, and the stimulus would appear to be much slower than the speeds obtained in the present study. The finding that the apparent speed is faster than its actual speed in the image plane is very much consistent with the previously noted findings of Clifford et al. (1999) and Bex and Makous (1997) who reported that stimuli simulating radially expanding optic flow appeared faster than their veridical speeds in the image plane.

Additionally, in conditions in which depth was given by stereomotion, the visual system does not rely on the *actual* interocular velocity difference (IOVD) between stereo images to derive perceived speed-in-depth. IOVD is essentially an image-plane cue, which can be used, in addition to the intended cue (i.e., changing disparity), to discriminate speed-in-depth (see, Brooks, 2002; Brooks & Stone, 2004; Fernandez & Farell, 2005; Harris & Watamaniuk, 1995). As mentioned, the stimulus used in the present study was a square plane of dots moving along the median depth plane towards the binoculus. To simulate this motion, monocular images consisted of dots that moved in opposite directions, but at the same two-dimensional speed. Thus, IOVD is indicative of a difference in the direction, but not speed, and under these conditions the predicted speed-in-depth agrees perfectly with the monocular speed of the dots (Brooks & Stone, 2004). Though this cue is present in our stimulus, as shown in Fig. 1, the observed perceived speed-in-depth is not consistent with the actual IOVD speed given by the dots, but it is in fact inferred as speed-in-depth, which is judged to be perceptually slower than a motion in the image plane that is physically matched in speed. These observations are in broad agreement with those of Brooks and Stone (2004) who showed that discrimination of speed-in-depth cannot be simply accounted for by inspection of the speed in each monocular image.

3. Experiment 2: the combination of image-plane speed and speed-in-depth produced by stereomotion and looming

Using the matched speeds obtained in Experiment 1, we examined how motion signals in the image plane and motion-in-depth are combined by the visual system to determine object speed. To address this issue, the same stimulus as in Experiment 1 was used, except that the square plane of dots always moved towards the observer with motion-in-depth defined either by stereomotion or by looming (Fig. 1B). Simultaneously with motion-in-depth, the same group of dots underwent motion in the image plane in random directions at a particular speed. This stimulus resembles a square traversing depth, with dots forming a dynamic swarming surface. Using this stimulus, we examined whether the apparent speed-in-depth of the square plane is affected by systematically changing the image-plane speed of the dots that form its surface.

3.1. Method

3.1.1. Observers

Four observers who participated in the previous experiment acted as observers in Experiment 2. All were naïve to the aims of the experiment.

3.1.2. Stimulus

The stereomotion stimulus was generated in a similar manner to that in Experiment 1, except that dots defining the square also underwent random-direction motion in the image plane. In addition, the dot-defined square was surrounded by a field of random dots ($6.94^\circ \times 9.25^\circ$) of identical dot density (18.70 dot deg^{-2}) moving at the same speed in the image plane as the dots forming the surface of the square plane. This procedure was implemented to ensure that any change in perceived speed-in-depth of this stimulus was a true reflection of an interaction between its speed-in-depth and the added random image-plane motion. As mentioned, to simulate motion towards the binoculus, dots in each monocular image moved in different directions, but at the same speed. It is possible that in each monocular image, the addition of random motion in the image plane (to the square) would change the perceived speed of the square (used to induce depth). This possibility exists since Khuu and Badcock (2002) demonstrated that the perceived speed of a cloud of dots is influenced by random 2D speed contained within the moving cloud. Therefore, any noted change in perceived speed-in-depth by the observer may be a consequence of this monocular interaction, and not between the actual speed-in-depth of the stimulus and the added speed in the image plane. It is important to note that this problem stems from an interaction between the IOVD speed of the stimulus and the added 2D random motion, and therefore a solution is to minimize IOVD and derive speed-in-depth exclusively through changing disparity alone. As mentioned above, this result can be achieved by placing the square in a field of 2D randomly moving dots, which effectively mask its boundaries. Thus, the stimulus is not appreciably visible in each monocular image, and the added random-direction motion in the image plane cannot influence the apparent speed of the translating square. However, as intended, the stimulus is visible through fusing the two monocular images, and its motion is defined by changing disparity (since this cue is not affected by the added random motion in the image plane). In the looming condition, the stimulus was constructed in the same manner as in Experiment 1, and additionally, random-direction motion was introduced to the dots that defined the square. These procedures resulted in the vivid percept of a looming stimulus with a noisy surface; the stimulus did not appear transparent.

3.2. Procedure

Experiment 2 examined the effect of changing 2D speed on the perception of speed-in-depth. Observers were presented pairs of motion-in-depth stimuli in a two-interval forced-choice task. In one interval, observers were presented with a reference stimulus, which moved in depth at one of the three speeds scaled individually according to the results of Experiment 1; with stereomotion: 1.51 cm/s, 2.10 cm/s, and 3.25 cm/s, for AL; 1.37 cm/s, 2.03 cm/s, and 2.69 cm/s, for TL, 1.44 cm/s, 2.03 cm/s, and 2.69 cm/s, for IT, and 1.63 cm/s, 2.34 cm/s, and 3.127 cm/s for NY; with looming: 1.73 cm/s, 2.17 cm/s, and 2.90 cm/s for AL, 1.03 cm/s, 1.78 cm/s, and 2.90 cm/s for BI, 1.66 cm/s, 2.30 cm/s, and 3.51 cm/s for JC, and 1.84 cm/s, 2.56 cm/s, and 4.23 cm/s for NY. Scaled values were individually derived in accordance with linear equation fits (deriving the slope and y-intercept) to data obtained in Experiment 1 for each observer. Importantly, this scaling ensured that for each observer any integration of image-plane speed and speed-in-depth reflects a combination of perceptually equal quantities. In addition to motion-in-depth, dots forming the surface of the stimulus moved in random directions at image-plane speeds of 0 cm/s, 0.40 cm/s, 0.80 cm/s, or 1.59 cm/s. As mentioned above, the random-direction motion in the image plane of dots does not contribute to stereomotion or looming, as there is no net global motion direction, but the moving dots provide an independent source of speed signals in the image plane that may influence the apparent speed-in-depth of the stimulus. In the other interval, the test stimulus, which was similar to the reference stimulus, was presented moving in depth (cued by stereomotion or looming depending on the condition) at a speed randomly chosen, but dots did not undergo random motion in the image plane. Therefore, the only feature distinguishing the reference and test stimuli is that the former contained dots undergoing random-direction motion in the image plane at a particular speed. The task of observers was to adjust the physical parameter that generated speed-in-depth of the test stimulus until its apparent speed matched the apparent speed-in-depth of the reference stimulus. As in Experiment 1, a trial consisted of a series of discrete presentations in which the test stimulus speed was adjusted from presentation to presentation, while the reference stimulus speed was fixed. Observers ended the experiment when they obtained a perceptual speed match between the two stimuli. A block of trials consisted of 24 trials: three levels of speed-in-depth and four levels of speed in the image plane which was repeated for both stereomotion and looming conditions. Observers performed these trials in a random order within a block of trials. A block was repeated 10 times and the results were averaged over blocks.

3.3. Results and discussion

Two outcomes are possible for Experiment 2. First, for the types of motion-in-depth cues used, it is not mandatory that speed in the image plane be combined with speed-in-depth. For this outcome, the apparent speed-in-depth would be unaffected by changing the image-plane speed of dots, and would implicate independent processes involved in the analysis of these two different motions. Second, a common mechanism may integrate speed in the image plane and speed-in-depth, and then apparent speed-in-depth would be affected by changing the image-plane speed of dots. Given this outcome, it would be possible to discern the nature of the analyses and whether they are similar for looming and stereomotion. In Figs. 3 and 4 the adjusted speed-in-depth of the test stimulus, scaled using the results of Experiment 1 to perceptually equate speed-in-depth and speed in the image plane, is plotted as a function of the speed of random-direction motion of dots in the image plane. The dotted line indicates the physical speed-in-depth of the stimulus, while separate columns illustrate data from stimuli with different speeds in depth; each row shows the results of each observer. Note that for each panel, the scaled speed in the image plane is given for each observer. The pattern of results is very similar for the four observers, and there are three notable findings from these data.

First, when the test stimulus contained no random-direction image-plane motion, and was therefore identical to the reference stimulus, observers were accurate in judging speed-in-depth, and corresponding speed adjustments of the reference stimulus, denoted by open white square symbols in Figs. 3 and 4, are close to dotted lines representing the actual speed-in-depth of the stimulus. Second, for stereomotion (Fig. 3), with the introduction of random-direction motion in the image plane to the reference stimulus, observers tended to judge its speed-in-depth to be faster than the actual value (solid black square symbols), and this effect increased as a function of the speed of random-direction motion in the image plane. It is important to note that for the image-plane speeds of 0.4 and 0.8 cm/s there are individual variations in the extent to which judged speed-in-depth was different from the actual speed of the reference stimulus, though a comparison between these two points shows an upwards trend (a possible explanation for this data trend is given below). However, for the fastest random-direction speed (1.59 cm/s) used in the present study, the apparent speed-in-depth of the stimulus is approximately 20% faster than its actual speed. This observation with stereomotion is unlikely to be a direct consequence of integrating motion signals in the image plane present in monocular images or derived from IOVD. As noted, the use of a random motion background effectively minimized IOVD, and while attentional tracking of dots could enhance monocular motion, since early motion detectors need only about 100 ms of stimulation to reach their peak response, it is dif-

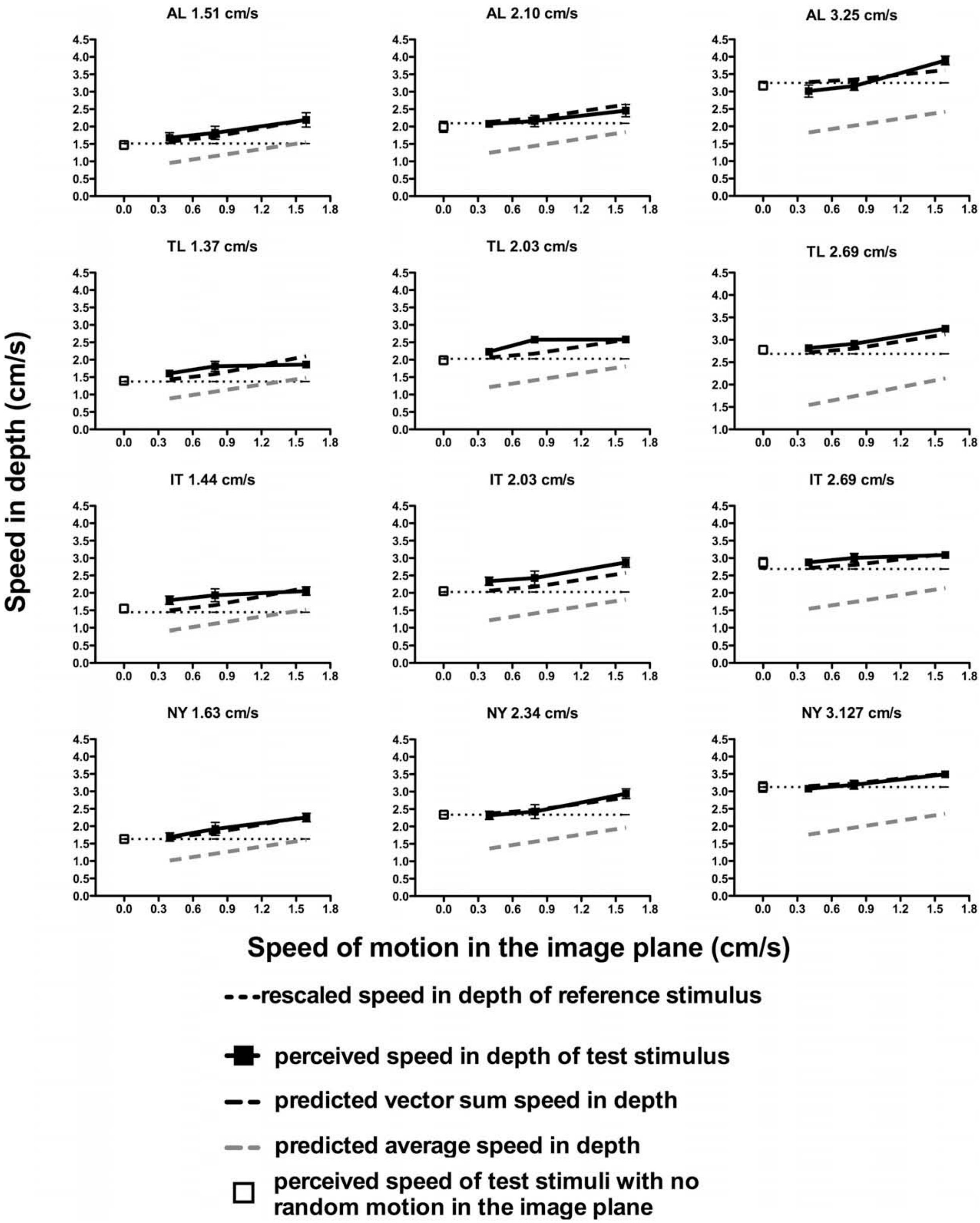

Fig. 3. The perceived speed-in-depth of the test stimulus, defined by stereomotion, is plotted against the speed of random-direction motion in the image plane (solid black squares and black solid line). The *y*-axes are scaled according to the ratios between speed in the image plane and speed-in-depth obtained from individual observers in Experiment 1. The dotted line represents the calculated physical speed-in-depth of the square. Open white squares denote the adjustment made to the test stimulus without random-direction motion in the image plane, which serves as the baseline to which judged speeds should conform if random-direction dot motion does not affect apparent speed-in-depth. Black and gray dashed lines correspond to predictions of the models of vector sum and of a simple average, respectively. Error bars represent ±1 SEM.

ficult to explain the present finding by any account that is not solely based on motion-in-depth inferred from motion signals in the image plane. Though, it is possible that attentional tracking could provide an effective monocular signal on which a reliable estimate of motion-in-depth is inferred. As shown in Fig. 3, observers did not match speed-in-depth with the random-direction image-plane dot speed present in the monocular images. Such an outcome would lead to a perceptual lowering of apparent speed-in-depth. Third, for looming, while apparent speed-in-depth also increases as a function of the speed of random-direction motion in the image

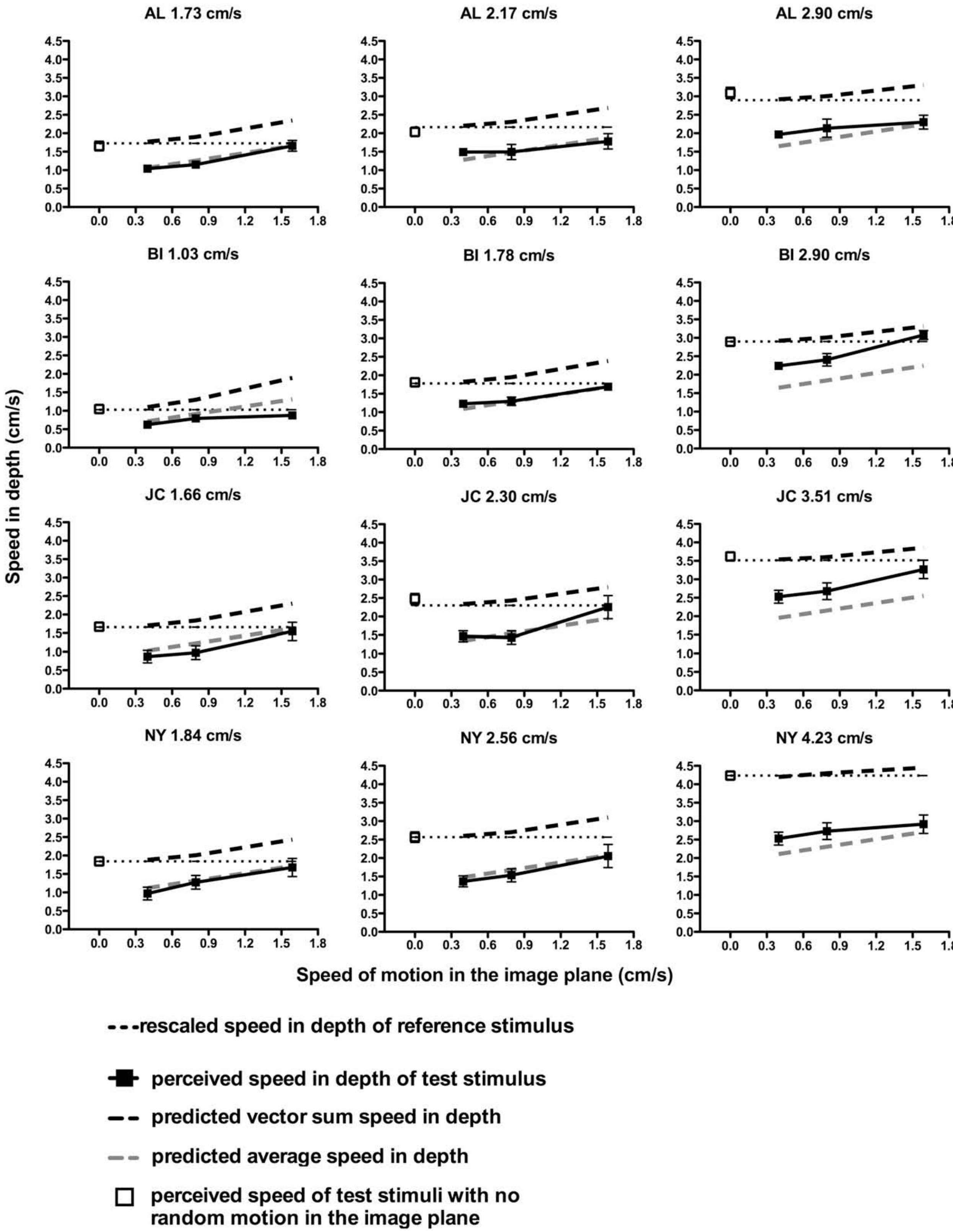

Fig. 4. The perceived speed-in-depth of the square, defined by looming, is plotted against the speed of the random-direction motion in the image plane (solid black squares and black solid line). The *y*-axes are scaled according to the ratios between image-plane dot speed and speed-in-depth obtained from individual observers in Experiment 1. The dashed line represents the calculated physical speed-in-depth of the square. Open white squares denote the adjustment without random-direction image-plane dot motion, which serves as a baseline. Black and gray dashed lines correspond to predictions of the models of vector sum and simple average, respectively. Error bars represent ±1 SEM.

plane, speed adjustments (solid black square symbols) for the range of random-direction speeds in the image plane tested tended to be lower than the actual speed-in-depth (dotted lines), as shown in Fig. 4. This decrease in perceived speed-in-depth is most evident for the image-plane speeds of 0.4 cm/s and 0.8 cm/s, while for the fastest image-plane speed of 1.59 cm/s judged perceived speed-in-depth resembles the actual test stimulus.

The very different set of results obtained with stereomotion and looming challenges the view that the signals derived from these two motion-in-depth cues are qualitatively similar and feed into an early common motion-in-depth mechanism (Regan & Beverley, 1979). Here, we demonstrate that under certain stimulus conditions, looming and stereomotion produce fundamentally different visual-system responses.

Two plausible solutions may provide an account for the pattern of results found in Experiment 2, which was concerned with how the visual system integrates speed in the image plane and speed-in-depth. One possible solution (though many have been proposed for the integration of local image-plane velocities; see, e.g., Bowns, 2006) is that the visual system initially extracts speed signals in the image plane and speed-in-depth signals and combines these estimates. Their independent analysis may arise because velocities in the image plane and in depth are physically orthogonal, and thus a vector-sum operation may be applied to provide an estimate of the resultant speed. The dashed black line in Figs. 3 and 4 represents the solution derived from this operation. An alternative solution is that the estimates of speed in the image plane and speed-in-depth are extracted jointly by the same mechanism, and they are simply averaged to provide an overall speed estimate, which is given by the dashed gray line in Figs. 3 and 4. This solution is plausible since the visual system has been shown to average local motion in the image plane in the perception of global speed in the image plane (Khuu & Badcock, 2002; Ross, 2004; Watamaniuk & Duchon, 1992) and is very much consistent with previous work examining the apparent speed of optic flow (Khuu & Badcock, 2002).

As illustrated in Fig. 3, a vector-sum operation is adequate to account for speed adjustments when stereomotion is used to define motion-in-depth. Consistent with our data, a vector sum solution predicts a small change in apparent speed-in-depth for the image-plane speeds of 0.4 cm/s and 0.8 cm/s, but a large increase for a speed of 1.59 cm/s. However, as shown in Fig. 4, a model of simple averaging provides a better account for the speed adjustment when motion-in-depth is defined by looming. Slower image-plane speeds (i.e., 0.4 cm/s and 0.8 cm/s) act to reduce the apparent speed of the test stimulus, but as previously mentioned, for the fastest image-plane speed used in the study, the perceived speed-in-depth is very similar to the actual speed of the reference stimulus (dashed line). This is not because there is no interaction between the two speed signals. Rather, it is because the image-plane speed is very similar to the actual speed-in-depth (with some variation between observers), and therefore their average is only minimally different to the component speeds.

There is a notable trend in Fig. 4 – for the fastest speed-in-depth, simple averaging consistently overestimates the reduction in apparent speed-in-depth when compared to the behavioral data. A key characteristic of the averaging solution employed in the present study is that it assumes equal weighting of speed in the image plane and speed-in-depth estimates regardless of their difference in magnitude. However, this assumption may not be valid, given previous research showing that speed in the image plane is processed by at least two broadly overlapping speed systems sensitive to low and high speeds (Edwards, Badcock, & Smith, 1998; Khuu & Badcock, 2002). It is possible that the overestimation of the averaging solution noted in Fig. 4 is due to relatively slow speeds in the image plane that may not be equally averaged with the fastest speed-in-depth. This overestimation is, perhaps, reflective of a process of weighted averaging, as the difference between the magnitudes of speed in the image plane and speed-in-depth is sufficiently large to activate different speed systems. This finding is very much consistent with the findings of Khuu and Badcock (2002) where they demonstrated that the apparent speed of a radially expanding flow pattern, divided into sectors and in which alternating sectors contained dots moving at different speeds, is equal to the average speed between sectors. However, Khuu and Badcock (2002) showed that averaging did not occur if sectors contained dots moving at speeds that activated separate systems. Additionally, an explanation for this dissociation between slow and fast speed processing is provided by the findings of Regan and Beverley (1979), who reported that looming is an effective cue to depth at slow speeds. The fastest speed employed in the present study may produce a weaker looming cue and it is likely that under these circumstances the visual system relies on the actual image-plane dot speed, rather than the interpreted speed-in-depth signaled by looming.

4. General discussion

The results of Experiment 2 lend themselves to the conclusion that different computational steps are employed by the visual system for the extraction of speed-in-depth when signaled by looming, than when signaled by stereomotion. This result is noteworthy since in natural scenes looming and stereomotion are inevitably associated, and coupled processing of these two cues has been reported for a variety of psychophysical judgments. Dissociation in the processing of looming and stereomotion is reported in the present study. It may be that the difference in processing found here is a consequence of the visual system representing motion-in-depth, signaled by stereomotion and by looming, at different stages of processing. A current consensus is that motion in the image plane is processed in at least two computational stages, where estimates of local velocities are initially extracted and are used to derive a global representation of motion across the image plane (e.g., Adelson & Movshon, 1982; Smith, Snowden, & Milne, 1994; Wilson, Ferrera, & Yo, 1992). Stereomotion can be effectively used to individuate the motion-in-depth of local elements. Looming, by contrast (and like optic-flow), cannot be employed in this fashion as motion-in-depth is generated by a particular global configuration of local image-plane velocities; local analyses of a looming pattern only reveals the image-plane trajectories of local elements, which is inadequate to produce an estimate of motion-in-depth.

The difference between stereomotion and looming has implications for an explanation of the findings of the present study. The major experimental manipulation of Experiment 2 was the introduction of speed signals in the image plane to determine their effect on the perception of speed-in-depth signaled by stereomotion or by looming. In the stereomotion condition, independent estimation of speed in the image plane and speed-in-depth at the local stage is possible, since the components of motion-in-depth, and random-direction motion in the image plane, introduced to an individual dot, are distinguishable. We demonstrate that the two estimates are combined by vector summation, which reflects an orthogonal combination of speed estimates in the image plane and in depth. This procedure is not viable in the looming condition, where information about the motion-in-depth of individual dots is unavailable. The looming stimulus, when considered locally, is composed of image-plane dot movements that are hardly distinguishable from the random-direction image-plane dot movements introduced to individual dots, which evidently leads the visual system to average the two velocity components, as in the determination of global speed in the image plane (see also, Khuu & Badcock, 2002; Watamaniuk & Duchon, 1992).

While the vector-sum model provides an account of the apparent speed of the stimulus, it should be noted that it predicts that the stimulus will follow an oblique trajectory in three-dimensional space which is inconsistent with the apparent object motion: a square plane of dots moving towards the observer in depth. The discrepancy between the predicted and perceived direction of trajectory traversed by the stimulus suggests that the speed and direction estimates of object motion are derived independently in three-dimensional space using different computational procedures. This result is consistent with a few studies indicating that speed and direction of motion are derived by separate mechanisms (Masson, Mestre, & Stone, 1999; Matthews, Luber, Qian, & Lisanby, 2001). Alternatively, vector summation may occur at the local level for each dot (thus dots may move along oblique trajectories), but given that image-plane dot-velocities are random, when the visual system combines local estimates to determine the global direction, there is no net change.

We conclude that the visual system employs different computational procedures in the extraction and combination of image speed in the image plane and in depth. The different results for stereomotion and looming are reflective of the differences by which the visual system extracts these depth cues.

Acknowledgements

Grateful acknowledgements are given to two anonymous reviewers and to the observers who participated in the study. S. Khuu is supported by funding from the Science Faculty Research Program, and an Early Career Research Award, from the University of New South Wales. A. Hayes was supported by a grant from Science Foundation Ireland.

References

Adelson, E., & Movshon, J. (1982). Phenomenal coherence of moving visual patterns. *Nature, 300*, 523–525.

Bex, P. J., & Makous, W. (1997). Radial motion looks faster. *Vision Research, 37*, 3399–3405.

Bowns, L. (2006). 'Squaring' is better at predicting plaid motion than the vector average or intersection of constraints. *Perception, 35*, 469–481.

Brenner, E., van den Berg, A. V., & van Damme, W. J. (1996). Perceived motion in depth. *Vision Research, 36*, 699–706.

Brooks, K. R. (2002). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision, 2*, 218–231. doi:10.1167/2.3.2. <http://journalofvision.org/2/3/2/>.

Brooks, K. R., & Stone, L. S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision, 4*, 1061–1079. doi:10.1167/4.12.6. <http://journalofvision.org/4/12/6/>.

Clifford, C. W. G., Beardsley, S. A., & Vaina, L. M. (1999). The perception and discrimination of speed in complex motion. *Vision Research, 39*, 2213–2227.

Edwards, M., Badcock, D. R., & Smith, A. T. (1998). Independent speed-tuned global-motion systems. *Vision Research, 38*, 1573–1580.

Fernandez, J. M., & Farell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research, 45*, 2786–2798.

Gray, R., Macuga, K., & Regan, D. (2004). Long range interactions between object-motion and self-motion in the perception of movement in depth. *Vision Research, 44*, 179–195.

Gray, R., & Regan, D. (1998). Accuracy of estimating time to collision using binocular and monocular information. *Vision Research, 38*, 499–512.

Harris, J. M., & Watamaniuk, S. N. (1995). Speed discrimination of motion-in-depth using binocular cues. *Vision Research, 35*, 885–896.

Heuer, H. (1993). Estimates of time to contact based on changing size and changing target vergence. *Perception, 22*, 549–563.

Hogervorst, M. A., & Brenner, E. (2004). Combining cues while avoiding perceptual conflicts. *Perception, 33*, 1155–1172.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford: Oxford University Press.

Khuu, S. K., & Badcock, D. R. (2002). Global speed processing: Evidence for local averaging within, but not across two speed ranges. *Vision Research, 42*, 3031–3042.

Ledgeway, T. (1999). Discrimination of the speed and direction of global second-order motion in stochastic displays. *Vision Research, 39*, 3710–3720.

Masson, G. S., Mestre, D. R., & Stone, L. S. (1999). Speed tuning of motion segmentation and discrimination. *Vision Research, 39*, 4299–4308.

Matthews, N., Luber, B., Qian, N., & Lisanby, S. H. (2001). Transcranial magnetic stimulation differentially affects speed and direction judgments. *Experimental Brain Research, 140*, 397–406.

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision Research, 21*, 491–500.

Poggio, G. F., Gonzalez, F., & Krause, F. (1988). The relationship between absolute disparity and ocular vergence. *Journal of Neuroscience, 8*, 4531–4550.

Regan, D., & Beverley, K. I. (1979). Binocular and monocular stimuli for motion in depth: Changing-disparity and changing-size feed the same motion-in-depth stage. *Vision Research, 19*, 1331–1342.

Regan, D., & Beverley, K. I. (1983). Visual fields for frontal plane motion and for changing size. *Vision Research, 23*, 673–676.

Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology of Visual Science, 27*, 584–597.

Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences, 4*, 99–107.

Reichardt, W. (1961). Autocorrelation, a principle for relative movement discrimination by the central nervous system. In W. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). New York: MIT Press.

Richards, W. (1977). Selective stereoblindness. In H. Spekreijse & L. H. van der Tweel (Eds.), *Spatial contrast: Report of a workshop* (pp. 109–115). Amsterdam: North-Holland.

Richards, W., & Regan, D. (1973). A stereo field map with implications for disparity processing. *Investigative Ophthalmology of Visual Science, 12*, 904–909.

Ross, J. (2004). The perceived direction and speed of global motion in glass pattern sequences. *Vision Research, 44*, 441–448.

Rushton, S. K., & Wann, J. P. (1999). Weighted combination of size and disparity: A computational model for timing a ball catch. *Nature Neuroscience, 2*, 186–190.

Savelsbergh, G. J., Whiting, H. T., & Bootsma, R. J. (1991). Grasping tau. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 315–322.

Smith, A. T., Snowden, R. J., & Milne, A. B. (1994). Is global motion really based on spatial integration of local motion signals? *Vision Research, 34*, 2425–2430.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32*, 931–941.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.